Modeling the Response of the Mediterranean Fruit Fly (Diptera:Tephritidae) to Cold Treatment

MARK R. POWELL

U.S. Department of Agriculture, Office of Risk Assessment and Cost Benefit Analysis, Washington, DC 20250

J. Econ. Entomol. 96(2): 300–310 (2003)

ABSTRACT Recent interceptions of live Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), larvae in fruit that had been cold-treated during transit from abroad led to a reevaluation of the scientific basis for the relevant regulatory treatment schedules. A time–temperature response surface model based on the original experimental data from 1916 was developed and evaluated based on subsequent experimental trials and recent surveillance data collected from shipping operations. The resultant model is reasonably robust and supports the conclusion that the previous treatment schedule falls short of the intended probit nine level of security. Given the vintage of the data, methodological inconsistencies among studies, and the potential consequences of new introductions, additional research is warranted. Quantitative analysis of the currently available data suggests that future studies regarding the efficacy of cold storage should focus on low-temperature, short-duration treatments, where uncertainty about performance appears greatest. The analysis of subsequent experiments also demonstrates that for cold treatment trials most often resulting in zero survivors, Bayesian statistical methods applied to a series of replicated trials of more manageable size offers a feasible alternative to conducting impracticably large trials.

KEY WORDS *Ceratitis capitata*, cold treatment, model

IN RESPONSE TO INTERCEPTIONS of live Mediterranean fruit fly (*Ceratitis capitata* (Wiedemann)) larvae in Clementines (several varieties of *Citrus reticulata*) that had been cold-treated during transit from Spain to the United States in 2001, a panel of United States Department of Agriculture (USDA) scientists and regulatory personnel was convened to review the available scientific literature and observations to date regarding the efficacy of the USDA Animal and Plant Health Inspection Service (APHIS) regulatory cold treatment schedule. Based on their review, the panel concluded that the previous T107-a treatment schedule falls short of the intended probit nine level of phytosanitary security (APHIS 2002a). The APHIS Plant Pest and Quarantine (APHIS/PPQ) Treatment Manual specifies the procedures and treatments (chemical and nonchemical) authorized to prevent the movement of plant pests into or within the United States (APHIS 2002b). The T107 series covers cold treatment of fruit. The required cold treatment is pest-, commodity-, and country-specific. APHIS cold treatment schedules T107-a, -c, and -f are authorized for control of *C. capitata* in various fruits. The probit nine level of security corresponds to a 3.2×10^{-5} probability of survival.)

The panel recommended increasing the length of cold treatment previously required at each temperature by 2 d. The panel also recommended that USDA establish research plans to verify the proposed new cold treatment parameters (APHIS 2002a). The panel's recommended revisions to the cold treatment schedule were incorporated into a proposed rulemaking for the importation of clementines from Spain (APHIS 2002c). Based on the analysis contained herein, the probit nine level of security could not be confirmed for the proposed low-temperature, short-duration treatments. Consequently, the final, revised cold treatment schedule was limited to the proposed treatments of 14 d or more (Table 1; APHIS 2002d).

The purpose of this paper was to develop and evaluate a response surface model relating Mediterranean fruit fly larval survival to cold treatment time–temperature combinations based on available data. Unlike a basic linear regression model with a single predictor variable, a response surface model can be plotted in at least three dimensions, indicating the response of the dependent variable (e.g., *C. capitata* survival) as two or more independent variables (e.g., time and temperature) are varied. A simple model is developed on the basis of multiple logistic regression analysis of larval survival data reported by Back and Pemberton

The opinions expressed herein are the views of the author and do not necessarily reflect the official policy or position of the U.S. Department of Agriculture. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government.

Table 1. Previous T107-a vs proposed and final cold treatment schedules

Temp, °C (°F)	Time (d)		
	Previous T107-a	APHIS (2002c) (Proposed)	APHIS (2002d) (Final)
0.0 (32)	10	12	–
0.6 (33)	11	13	–
1.1 (34)	12	14	14
1.6 (35)	14	16	16
2.2 (36)	16	18	18

(1916), the original research, which informed development of the regulatory treatment schedule (T107) (APHIS 2002a). The predictions of the response surface model are then compared with results of subsequent *C. capitata* larvae cold treatment trials conducted under the relevant time–temperature treatment combinations (Nel 1936, Sproul 1976, Hill et al. 1988, Jessup et al. 1993, Santaballa et al. 1999, De Lima et al. 2002) as well as surveillance data collected from marine shipping vessels in 2001 (APHIS 2002e). The analysis is not intended to elaborate the definitive model of *C. capitata* larval response to cold treatment, rather it primarily aims to corroborate whether the previous cold treatment schedule fails to achieve the intended level of protection, assess broad trends in the available data (e.g., investigate the relative importance of cold treatment temperature and duration and the roles of uncertainty and variability in treatment efficacy), and provide input regarding the focus of future data acquisition.

Materials and Methods

Response Surface Model Development. Because of their coverage of time–temperature combinations spanning the entire range of concern to phytosanitary programs and the reporting of unsummarized results, Back and Pemberton (1916) provide the best available data for development of a response surface model to estimate *C. capitata* larval survival under cold treatment. Apples, peaches, and kamani nuts were used as *C. capitata* hosts in their studies. Six cold-storage temperature levels were included in the new analysis, with temperature converted to Celsius and coded as the midpoint in the case of nominal storage temperature intervals: 32°F (0°C), 32–33°F (0.28°C), 33–34°F (0.83°C), 34–36°F (1.67°C), 36°F (2.22°C), and 36–40°F (3.33°C). The final storage temperature level was included to inform the high-temperature, long-duration region of the response surface. (Because the 36°F (2.22°C) treatments reported by Back and Pemberton (1916) were limited to 16 d or less for larvae, evaluation of the proposed 18-d cold treatment (Table 1) otherwise would be based on extrapolation.) Data on exposures at the 38–40°F (3.33–4.44°C) and 40–45°F (4.44–7.22°C) storage temperature levels reported by Back and Pemberton (1916) were excluded to limit the effect of independent variable measurement error on the multiple regression analysis. The duration of cold storage varied by storage temperature, with a minimum of 15 d and a maximum of 30 d. Back and Pemberton (1916) reported unsummarized data on the number of larvae found alive and dead (separately) at each time–temperature combination.

Although Back and Pemberton (1916) did not completely report their methodology, it appears that the duration of treatment refers to cold storage time, instead of the cold treatment time elapsed once the fruit cooled to a given temperature. This presents a potentially significant source of measurement error and an inconsistency with more recent studies and the cold treatment regulatory requirements (T107). Mason and McBride (1934) reported that the time required for the interior of fruits (apples, oranges, and avocados) to reach storage room temperatures of 28–31°F (−2.22 to −0.56°C) ranged from 18 to 48 h. In general, precooling time depends on the volume and packing of fruit being treated. Conceivably, larval survival under cold treatment may depend on the cooling rate (i.e., a biological response may differ if stress is applied gradually or swiftly, depending on the organism's acclimation potential). Available *C. capitata* cold treatment studies also differ in the survival measurement endpoint. Back and Pemberton (1916) measured larval survival based on observation 24–48 h after removal from cold storage. Although the detection of live larvae may be judged sufficient for phytosanitary inspection purposes, some subsequent studies recorded survivors only as those that emerged from the fruit and attained the pupal stage because some larvae that appear alive after cold storage fail to pupate (Mason and McBride 1934). Furthermore, in infested fruits that are not subject to cold treatment, survival of pupae to the adult stage may vary from 70 to 80% (Santaballa et al. 1999). Consequently, the response surface model based on the Back and Pemberton (1916) cold storage data are hypothesized to demarcate a plausible upper-bound on effective larval survival under cold treatment conducted in compliance with regulatory requirements.

Back and Pemberton (1916) identified the observed larval stage as first, second, or third instars, and statistical analysis tends to support the conclusion that later instars are somewhat more cold-tolerant (see discussion below). Some subsequent studies failed to distinguish larval stage, however. For each time-temperature combination, therefore, Back and Pemberton (1916) data on first, second, and third instars were combined to develop the response surface model.

The time–temperature response surface was obtained using a standard logistic regression procedure (SAS PROC LOGISTIC), and assuming a simple main effects model:

$$\text{logit}(p_s) = \ln\left(\frac{p_s}{1-p_s}\right) = b_0 + b_1 * \text{temp}(^\circ C) + b_2 * \text{time}(d), \quad [1]$$

$$\text{where: est.prob.survival } (\hat{p}_s) = \frac{\exp(\text{logit}(p_s))}{1 + \exp(\text{logit}(p_s))}$$

$$= \frac{1}{1 + \exp(-(b_0 + b_1 * \text{temp}(^\circ C) + b_2 * \text{time}(d)))}$$

Table 2. Results of logistic regression on Black and Pemberton (1916) data

Link function	Model fit statistic (− 2log (lik))	
	Without log transformation	With log transformation
Logit	14279.725[a]	15644.654
Normit	14554.189	16656.654
Clog-log	16119.614	22878.081

Logit model parameter	Estimate	Standard error		$P > \chi^2$
		Binomial dispersion	Empirical dispersion	(Empirical dispersion)
b_0-Intercept	6.6448	0.0884	0.5234	<0.0001
b_1-Temp, °C	0.3063	0.0199	0.1177	0.0093
b_2-Time, d	−1.1155	0.0127	0.0753	<0.0001

[a] Best fit model: logit without log transformation.

Under this model, the logit link function is assumed to transform the underlying model into a linear function of the parameters (equation 1, and the error about the fit regression curve is assumed to be binomially distributed (Brown and Rothery 1993).

The regression model parameter estimates are obtained by maximizing the likelihood function (lik):

$$\text{lik} = \Pi_i \hat{p}_s^s * (1 - \hat{p}_s)^{n-s}, \qquad [2]$$

where s = survivors
n = no.larvae treated
i = ith treatment

Maximum likelihood estimates for the parameter values are commonly obtained by minimizing the −2log(lik) function, using initial parameter values. Note that this method avoids taking the logarithm of zero or dividing by zero (equation 1), such that data from trials with zero or 100% survivors do not generally present a problem encountered using weighted least squares or similar parameter estimation methods.

In developing the response surface model, three generalized linear model link functions were considered: the logit, normit, and complementary log-log (clog-log). Each was fitted with and without a logarithmic transformation of time and temperature. The logit is the inverse of the cumulative logistic distribution function. Like the normal distribution, it is symmetric about the mean, but the logistic is a more heavy-tailed distribution (i.e., the logit presumes greater variation in population response to a given treatment). The normit distribution is the inverse of the cumulative standard normal distribution function. The more familiar term probit is often used, although conventionally the probit function contains the additive constant five to avoid negative values. (Thus, the probit nine level refers to the area under the standard normal distribution beyond four standard deviations above the mean, or simply 3.2×10^{-5}.) Applying both a logarithmic transformation and the probit link function assumes the tolerances to be lognormally distributed within the population. The clog-log function is the inverse of the cumulative extreme-value function (also called the Gompertz distribution), which is skewed. Although these represent a range of model forms, they are empirical models, rather than theoretically based models that imply a biological understanding of the mechanism of larval mortality as a result of low temperature. Among the models considered, the model based on the untransformed data and the logit link function (equation 1 was selected on the basis of goodness-of-fit criteria. (The best fitting model has the minimum −2log(lik) value in Table 2.)

Using a more flexible empirical model with additional terms would improve the statistical goodness-of-fit, but the candidate models were selected on the basis of parsimony and ease of interpretation. Including a temperature × time interaction term in the regression model was rejected because it produced nonsensical results, i.e., the model predicted higher survival at lower temperatures for treatment durations of 9 d or more. Independent variable measurement error (as a result of cooling time and broad cold treatment temperature intervals) may have contributed to these spurious results. Statistical diagnostics also identified several influential data points that represent potential outliers. The vintage of the Back and Pemberton (1916) dataset, however, precludes a reliable means of identifying and eliminating bona fide outliers from the analysis.

Results and Discussion

Regression Analysis Results. The multiple logistic regression analysis results presented in Table 2 indicate that the model parameters maximum likelihood estimates are statistically significant ($P < 0.01$). The markedly different variance estimates for the model parameters (depending on whether the error is assumed to be binomially distributed or is estimated empirically by dividing the deviance goodness-of-fit statistic by its degrees of freedom) does not affect the parameter estimates but indicates substantial extra-binomial dispersion. (The error distribution is broader than under the binomial assumption.) Fig. 1 presents the response surface model. The lower confidence bound for 32°F (0°C) and the upper confidence bound for 36°F (2.22°C) are presented as the outermost solid lines and are indicated by their respective symbols. The confidence bounds account for parameter uncertainty and are derived from the variance-covariance matrix obtained while relaxing the binomial error distribution assumption. Although temperature was

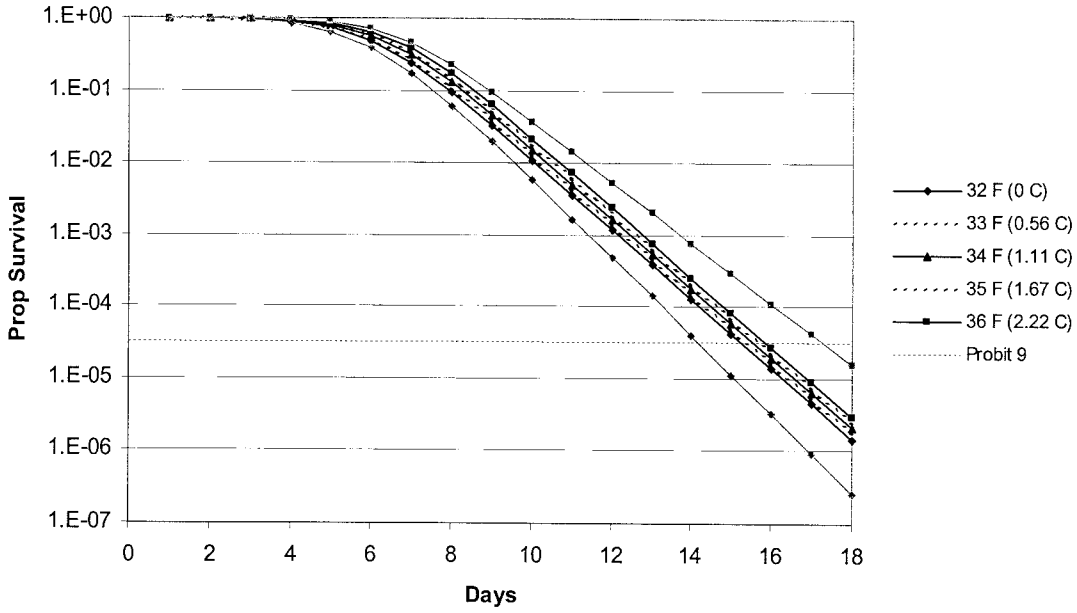

Fig. 1. Response surface model for *C. capitata* larval survival under cold storage

found to be statistically significant, the model suggests that within the range of cold storage conditions considered, one additional day of cold storage may yield substantially more protection than lowering the storage temperature by 1°F (0.56°C).

Given the vintage of the data and methodological concerns, these results should only be considered as a hypothesis to be evaluated through additional data acquisition and analysis. Nevertheless, for cold treatment periods of <16 d, the results tend to support the finding that the previous T107-a treatment schedule typically achieves less than the probit nine level of phytosanitary security ($p_s = 3.2 \times 10^{-5}$), as well as the conclusion that the duration of cold treatment would need to be increased if this level of security is to be achieved (APHIS 2002a).

Model Evaluation. To permit an assessment of the fit of the response surface model to the Back and Pemberton (1916) data, Fig. 2 overlays the data with the model predictions and extra-binomial 95% confidence bounds, all expressed on a logit scale. Presentation of the reported data on the linearizing logit scale requires some form of statistical treatment of trials with zero or 100% survivors to avoid taking the log of zero or dividing by zero. Therefore, such trials are indicated in Fig. 2 by open symbols and are represented by the value $(s \pm 0.5)/(n+1)$, where s is the number of survivors and n is the number of larvae treated. Note that this treatment is merely a common statistical convention used for the purpose of graphic presentation and does not affect the regression analysis. These data points are highlighted because the apparent differences in survivorship are more a function of sample size than biology. In assessing the fit of the model to the Back and Pemberton (1916) data, note that the confidence bounds in Fig. 2 represent uncertainty about the true mean response only (i.e., logit model parameter uncertainty) and do not account for random variability about the mean response. (Only approximate methods are available to estimate prediction intervals for nonlinear models.) As indicated above, a more flexible empirical model would improve the fit of the model to the data.

Although goodness-of-fit between a model and the data on which it is based is an important model evaluation criterion, the robustness of the model to independent data may be more important for the purposes of testing the conclusions drawn from it. To assess the robustness of the model based on the Back and Pemberton (1916) data, model predictions were compared with 95% confidence intervals constructed about the results of *C. capitata* larvae cold treatment trials conducted under similar time–temperature combinations, as well as recent surveillance of shipping operations. As indicated above, however, some subsequent studies failed to report the insect stage. Therefore the model was first evaluated regarding sensitivity to the effect of insect stage.

Multiple logistic regression analysis of the Back and Pemberton (1916) data with a model including categorical variables for insect stage:

$$\text{logit}(p_s) = f(temp, time, stage) \qquad [3]$$

indicates that eggs were less likely to survive a given time–temperature combination than larvae. Also, first and second instars were less likely to survive a given time–temperature combination than third instars (Table 3). In comparing the probability of a binary outcome for group A versus group B, an odds ratio of 1.0 indicates a lack of association between the independent variable and the dependent variable. If the odds ratio is <1.0, group A is less likely to have the outcome

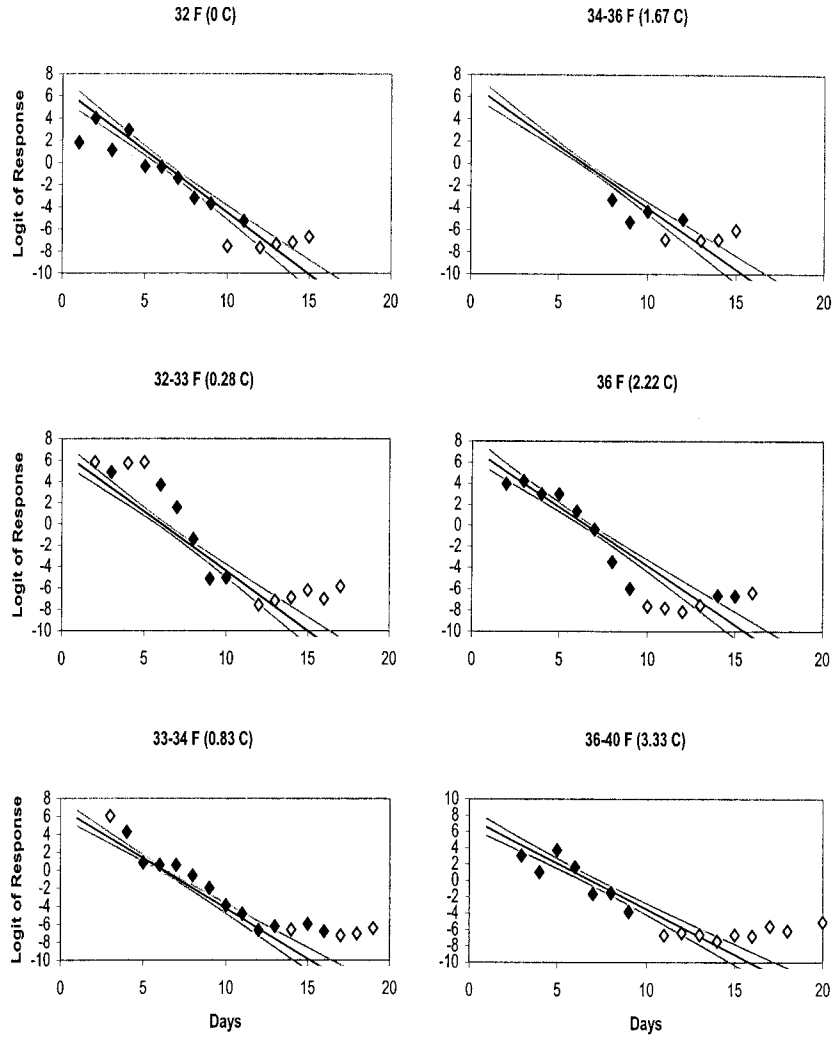

Fig. 2. Back and Pemberton (1916) data, with model fit and confidence bounds

than group B. In the comparisons, note that the 95% CL about the odds ratios do not contain 1.0 (Table 3). Because of a large proportion of incomplete cases, the eggs versus larvae comparison presented in Table 3 was conducted separately from the instar comparisons. In an analysis of the reduced set of complete cases containing data on all four stages, only the eggs versus third-instar comparison was statistically significant (odds ratio = 0.657, 95% CL: 0.592–0.729).

Table 3. Survival odds ratios for *C. capitata* stages

Comparison	Estimate (est)	Standard error	Odds ratio	95% Wald confidence limits
Eggs vs. larvae	−0.4825	0.0385	0.617	0.572–0.666
Instar 1 vs. 3	−0.4570	0.0589	0.633	0.564–0.711
Instar 2 vs. 3	−0.1357	0.0561	0.873	0.782–0.975

Odds ratio = exp(est), 95% Wald confidence limits = exp(est ± 1.96*Std. error).

Despite the statistically significant odds ratio comparisons, the magnitude of the effect of larval stage on the cold treatment time estimated to achieve a probit nine level of security appears insubstantial. Figure 3, for example, overlays the Back and Pemberton (1916) instar-specific data reported for cold treatment at 32°F (0°C) with the corresponding model predictions. (Because of the log scale on the y-axis, trials with zero survivors are indicated in Fig. 3 by open symbols and are represented by the value 0.5/(n+1), where n is the number of larvae treated.) Similarly, Jessup et al. (1993) and Santaballa et al. (1999) reported overlap among instars in the 95% confidence intervals for the time required to achieve a given level of mortality.

Various immature stages of fruit fly may be present in a given commercial consignment of fruit (Sproul 1976, Santaballa et al. 1999). To be reliable, therefore, cold treatments must be designed for the most cold-tolerant stage of *C. capitata.* Despite the apparent lack of sensitivity of modeled results to instar stage, in cases

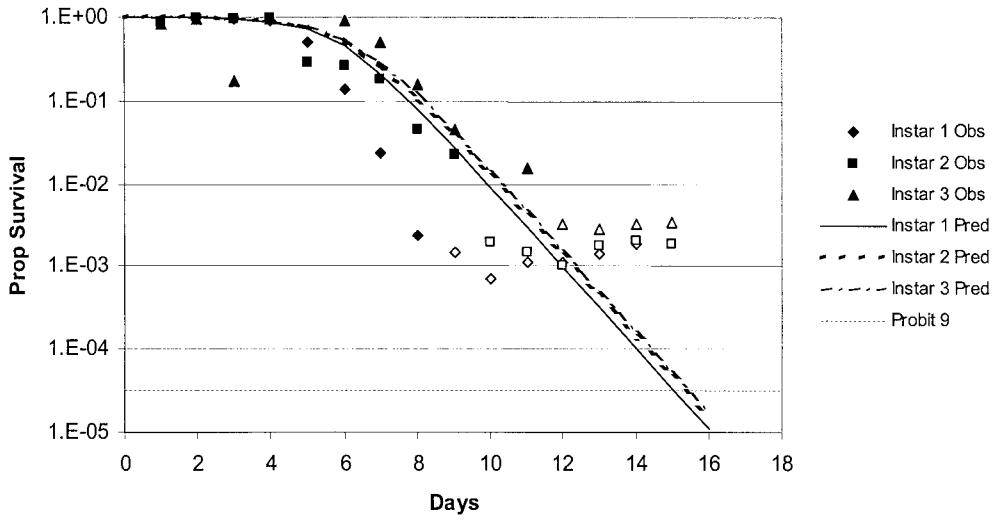

Fig. 3. *C. capitata* instar survival at 32°F (0°C)

where subsequent studies distinguished among observed larval stages, the predictions of the response surface model were compared with confidence intervals for reports specific to more mature or cold-tolerant larvae. (In some studies, second instars were judged to be most cold tolerant.) Table 4 presents the data used to evaluate the time–temperature response surface model.

The fruits, *C. capitata* larval stages and strains, and survival measurement endpoints used differ among studies. Nel (1936) used nectarines, peaches, plums, and grapes as host material and used live larvae as the survival measurement endpoint. The *C. capitata* larval stage was unidentified. Sproul (1976) used Granny Smith apples, and the survival measurement endpoint was emergent pupae. The data presented are for old larvae, primarily third instars of *C. capitata.* Hill et al. (1988) used Valencia and Navel oranges, and the survival measurement endpoint was emergent pupae. The data presented are for old larvae, primarily third instars. Jessup et al. (1993) used Lisbon and Eureka lemons, and the survival measurement endpoint was emergent pupae. The data presented are for primarily second instars, which were found to be most cold tolerant. Santaballa et al. (1999) used clementines, and the survival measurement endpoint was live larvae. The data presented for trials conducted up to 14 d are for old larvae, primarily third instars; young larvae (first and second instars) were used for the 16-d trials. (In the former smaller trials, no statistically significant differences in cold tolerance were observed between young and old larvae.) De Lima et al. (2002) used Lisbon lemons for 16 d trials and Navel and Valencia oranges and Ellendale and Murcott tangors for 18 d trials. The survival measurement endpoint was emergent pupae, and the data presented are for primarily second instars. Finally, in surveillance conducted by APHIS during 2001, *C. capitata* larvae were detected

Table 4. Cold treatment data used to evaluate response surface model

Study	Temp, °C (°F)	Time (d)	No. larvae treated (n_i)	Survivors(s)
Nel (1936)	1.1 (34)	9	1,670	50
	1.1 (34)	10	1,140	26
	1.1 (34)	11	2,905	3
	1.1 (34)	12	2,902	0
	1.1 (34)	12	3,907	0
Sproul (1976)	0.5 (32.9)	14	22,000	0
	0.5 (32.9)	14	800	0
	0.5 (32.9)	14	10,100	0
	0.5 (32.9)	14	12,900	0
	0.5 (32.9)	14	13,700	0
	1.5 (34.7)	16	8,000	0
	1.5 (34.7)	16	12,000	0
	1.5 (34.7)	16	21,800	0
	1.5 (34.7)	16	13,200	0
Hill et al. (1988)	1 (33.8)	16	18,904	0
	1 (33.8)	16	11,668	0
	1 (33.8)	16	10,584	0
	1.5 (34.7)	16	41,099	3
Jessup et al. (1993)	1 (33.8)	14	10,010	0
	1 (33.8)	14	10,140	0
	1 (33.8)	14	10,080	0
	1 (33.8)	14	20,015	0
	1 (33.8)	14	13,158	0
	1 (33.8)	14	10,170	0
Santaballa et al. (1999)	2 (35.6)	10	935	10
	2 (35.6)	12	935	5
	2 (35.6)	14	935	0
	2 (35.6)	16	11,317	0
	2 (35.6)	16	10,295	0
	2 (35.6)	16	10,376	0
De Lima et al. (2002)	2 (35.6)	16	141,441	0
	2 (35.6)	16	165,894	0
	2 (35.6)	16	133,788	0
	2 (35.6)	16	108,732	0
	2 (35.6)	16	132,216	0
APHIS (2002e)	0 (32)*[a]	10*[a]	212	2

[a] Assumed.

in 29 of 20,460 clementines that were cut and inspected after cold-treatment on 80 shipping vessels. A total of 212 larvae were detected, 2 live and 210 dead (APHIS 2002e). Larval stage was unidentified. Cold treatments likely varied among the shipping vessels, but a treatment of 0°C (32°F) for 10 d is assumed because it is the shortest treatment compliant with the previous T107-a schedule, and because the commodity is highly perishable. (The minimum marine shipping vessel transit time between Spain and the New York metropolitan area, for example, is approximately 7 d (Shipguide.com 2002). Although many of the shipments associated with the interceptions during 2001 may have used longer (e.g., 11–12-d) cold treatments (APHIS 2002d), evaluation of the model with respect to the surveillance data are robust to departures from the assumed treatment time and temperature, as indicated below.) Although not subject to experimental controls, the 2001 surveillance data arguably represents the best available evidence regarding the operational performance of cold treatment.

Confidence intervals constructed about the experimental trial and surveillance results were obtained assuming that the probability of larval survival for a given time–temperature combination is beta distributed. Because the beta distribution is the conjugate prior for the proportion of successes (p_s) when values (s) from a sample (n) follow a binomial distribution, the beta distribution is used to characterize the uncertainty about proportions arising from binomial processes (Vose 2000). To estimate the beta distribution parameters for trials with some surviving *C. capitata* larvae, the method of matching moments was used (Evans et al. 1993):

$$\text{If } s \sim \text{binomial } (n,p_s), \qquad [4]$$

$$\text{where: } p(s = x) = C_x^n p^x (1 - p)^{n-x},$$

$$\text{then } p_s \sim \text{beta}(\alpha, \beta),$$

$$\text{where: } \hat{\alpha} = \bar{x}\{[\bar{x}(1 - \bar{x})/s.d.^2] - 1\},$$

$$\hat{\beta} = (1 - \bar{x})\{[\bar{x}(1 - \bar{x})/s.d.^2] - 1,$$

$$\bar{x} = \frac{\text{survivors } (s)}{\text{no. larvae treated } (n)},$$

$$s.d.^2 = \bar{x}(1 - \bar{x})/(n - 1).$$

For trials in which no larvae survived cold treatment, sample moments cannot be obtained. (Estimating the beta distribution parameters would involve dividing by zero, because the sample variance (SD^2) is zero.) For these trials, Bayesian statistical methods were used to estimate the beta distribution parameters (Vose 2000). Bayes Rule implies:

$$\text{posterior prob. } (p_s|s,n) \propto$$

$$\text{prior prob. } (p_s)*\text{lik}(s,n|p_s),$$

$$\text{where: lik } (s = 0,\ n = n_i|p_s) = (1 - p_s)^{n_i} \qquad [5]$$

Note that the likelihood of observing zero out of n_i survivors, given p_s, follows from the binomial distribution where $s = 0$ (equation 4). The Bayesian estimation operation proceeded in sequential fashion, beginning with a uniform prior distribution, evaluating the likelihood of observing the results of the first trial ($s = 0, n_1$) for 10,000 discretized p_s values (from 1×10^{-7} to 1×10^{-3} at increments of 1×10^{-7}). The uniform prior implies that all p_s values from 1×10^{-7} to 1×10^{-3} are considered equally likely before considering the data. A posterior uncertainty distribution is obtained by calculating the product of the initial prior and the likelihood of having observed the data over the values of p_s. The resultant posterior uncertainty distribution then becomes the prior distribution for the second trial ($s = 0, n_2$), and so on. The process is repeated until all of the trials from a given study with zero survivors for a particular time-temperature combination have been evaluated. The moments of the final posterior distribution can then be calculated to obtain beta distribution parameter estimates (equation 4). Given estimated beta distribution parameter values, 95% credible intervals for each of the relevant time-temperature combinations were constructed by obtaining the 2.5th and 97.5th percentiles of the distribution (using the Microsoft© Excel BETAINV function). The term credible interval is commonly used in Bayesian statistics as an analog to the conventional confidence interval. Below, the conventional terminology is used regardless of the estimation method. Table 5 contains a worked example of the Bayesian procedure calculations for three trials reported by Santaballa et al. (1999).

Figure 4 presents the confidence intervals constructed about the evaluation data overlaid with the response surface model based on the Back and Pemberton (1916) data. (The response surface model predictions are presented as alternating bands of solid and dashed lines, as in Fig. 1. Lower and upper confidence bounds for the response surface model are presented as the outermost solid lines for 32°F (0°C) and 36°F (2.22°C), respectively. The probit nine level is presented as a horizontal reference line.) Any statistical model, but especially one based on data that is >80 yr old, should be greeted with a healthy dose of skepticism. Nevertheless, the model appears reasonably robust, particularly as an upper-bound on *C. capitata* larval survival under cold treatment. Four observations from two subsequent studies (Hill et al. 1988, Santaballa et al. 1999) indicate a mean survival proportion above the predicted response, but the confidence limits contain the model prediction in each case. Note also the consistency between the APHIS 2001 surveillance data and the model prediction. Although the parameters of the cold treatments performed on the shipping vessels were assumed to be 32°F (0°C) for 10 d, the confidence interval about the observed proportion of larval survival overlaps with the model predictions assuming that the cold treatments varied from 32°F (0°C) for 9 d to 36°F (2.2°C) for 12.5 d.

In general, the evaluation data analysis and the model predictions are particularly consistent in the low-temperature, short-duration region of the re-

Table 5. Example of Bayesian procedure calculations

Santaballa et al. (1999) 2°C (35.6°F), 16 d		Trial 1			Trial 2			Trial 3				
	s	0		s	0		s	0				Mean
	n	11317		n	10295		n	10376				3.13E-05
p	unif. prior	lik	product1	posterior 1	lik	product2	posterior 2	lik	product3	posterior 3	(p-m) ^2	Var
1.00E-07	1.00E-04	9.99E-01	9.99E-05	1.13E-03	9.99E-01	1.13E-03	2.16E-03	9.99E-01	2.16E-03	3.19E-03	9.74E-10	9.77E-10
2.00E-07	1.00E-04	9.98E-01	9.98E-05	1.13E-03	9.98E-01	1.13E-03	2.15E-03	9.98E-01	2.15E-03	3.18E-03	9.68E-10	
3.00E-07	1.00E-04	9.97E-01	9.97E-05	1.13E-03	9.97E-01	1.13E-03	2.15E-03	9.97E-01	2.14E-03	3.17E-03	9.62E-10	alpha
4.00E-07	1.00E-04	9.95E-01	9.95E-05	1.13E-03	9.96E-01	1.12E-03	2.15E-03	9.96E-01	2.14E-03	3.16E-03	9.55E-10	1.003204
5.00E-07	1.00E-04	9.94E-01	9.94E-05	1.13E-03	9.95E-01	1.12E-03	2.14E-03	9.95E-01	2.13E-03	3.15E-03	9.49E-10	beta
6.00E-07	1.00E-04	9.93E-01	9.93E-05	1.12E-03	9.94E-01	1.12E-03	2.14E-03	9.94E-01	2.12E-03	3.14E-03	9.43E-10	32040.22
–	–	–	–	–	–	–	–	–	–	–	–	
9.995E-04	1.00E-04	1.22E-05	1.22E-09	1.38E-08	3.38E-05	4.66E-13	8.90E-13	3.12E-05	2.77E-17	4.11E-17	9.37E-07	2.5%ile
9.996E-04	1.00E-04	1.22E-05	1.22E-09	1.38E-08	3.38E-05	4.65E-13	8.88E-13	3.11E-05	2.76E-17	4.09E-17	9.38E-07	8.0E-07
9.997E-04	1.00E-04	1.21E-05	1.21E-09	1.37E-08	3.37E-05	4.64E-13	8.86E-13	3.11E-05	2.76E-17	4.08E-17	9.38E-07	50%ile
9.998E-04	1.00E-04	1.21E-05	1.21E-09	1.37E-08	3.37E-05	4.63E-13	8.84E-13	3.11E-05	2.75E-17	4.07E-17	9.38E-07	2.2E-05
9.999E-04	1.00E-04	1.21E-05	1.21E-09	1.37E-08	3.37E-05	4.62E-13	8.82E-13	3.10E-05	2.74E-17	4.05E-17	9.38E-07	97.5%ile
1.000E-03	1.00E-04	1.21E-05	1.21E-09	1.37E-08	3.36E-05	4.61E-13	8.80E-13	3.10E-05	2.73E-17	4.04E-17	9.38E-07	1.2E-04
sum	1.00	sum	0.088304		sum	0.523403		sum	0.675288			

lik = eq.5, product1 = unif.prior*lik, product2 = posterior1*lik, product3 = posterior2*lik, $\text{posterior}_i = \text{product}_i/\text{sum}(\text{product}_i)$, Mean(m) = Σposterior3*p, Var = Σposterior3*(p-m)^2, alpha, beta = eq.4, %ile = Betainv(%ile, alpha, beta)

sponse surface, and both lines of evidence suggest that previous cold treatment requirements fall short of achieving the probit nine level of security (3.2×10^{-5}). Specifically, the confidence interval about the 34°F (1.1°C), 12-d trial reported by Nel (1936) departs significantly from the response surface model, but both indicate a low level of confidence (<50%) that the previous T107-a combination of 34°F (1.1°C), 12 d achieves the probit nine level of security. In contrast, the confidence interval about the trial reported by Jessup et al. (1993) for 33.8°F (1°C), 14 d suggests a moderate degree of confidence (<95%) that the proposed 34°F (1.1°C), 14-d treatment (APHIS 2002a) should achieve the probit nine level of security. The response surface model and the confidence intervals about 3 of the 16 d trials (Sproul (1976) at 34.7°F (1.5°C), Santaballa et al. (1999) at 35.6°F (2°C), and De Lima et al. (2002) at 35.6°F (2°C)) suggest a moderate degree of confidence (<95%) that the proposed 35°F (1.6°C), 16-d treatment (APHIS 2002a) should achieve the probit nine level of security. Both the response surface model and the confidence interval about the 35.6°F (2°C), 18-d trial reported by De Lima et al. (2002) indicate a high degree of confidence (>95%) that the proposed 36°F (2.2°C), 18-d treatment (APHIS 2002a) should achieve the probit nine level of protection.

Although these results suggest that the 36°F (2.2°C), 18 d treatment may be more than sufficient to attain the probit nine level of security, note that the confidence interval about the 34.7°F (1.5°C), 16-d trial reported by Hill et al. (1988) indicates a low level of

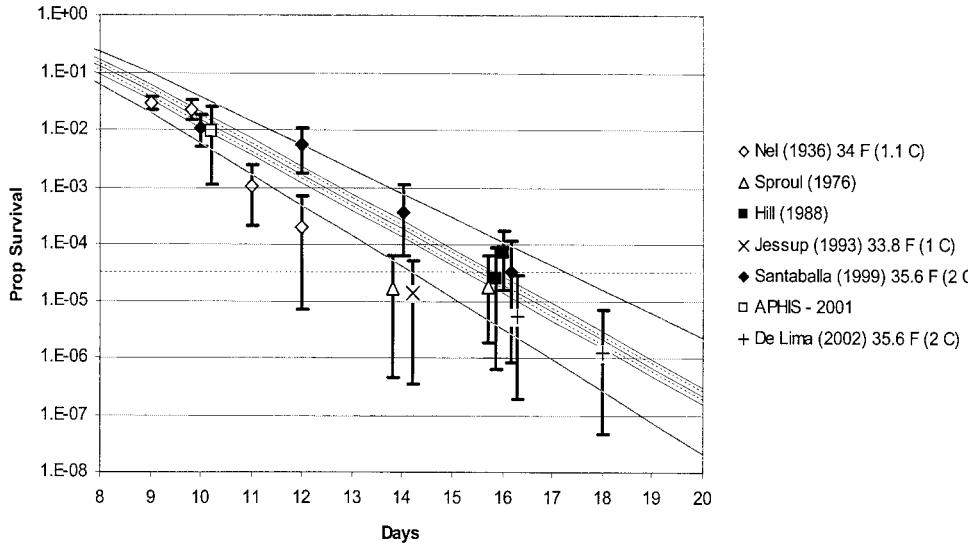

Fig. 4. Evaluation of response surface model for *C. capitata* larval survival under cold storage

confidence (<50%) of achieving the probit nine level. This serves as a reminder that some variability in cold treatment performance can be expected depending on cold treatment procedures (e.g., precooling time), study materials and methods, and other factors. Note, however, the confidence intervals for the 16-d trials conducted at the same temperatures but with different fruits (Sproul (1976) and Hill et al. (1988) at 34.7°F (1.5°C); Santaballa et al. (1999) and De Lima et al. (2002) at 35.6°F (2°C)). Here, the confidence intervals span multiple orders of magnitude, although the point estimates differ by less than an order or magnitude. This suggests that at very low levels of survival (ca. probit 9), the uncertainty associated with larval response to a specific time-temperature-fruit combination may be greater than the variability in response because of different fruit hosts, at least for the cultivars considered.

Uncertainty remains regarding what statistical model form best describes the observed cold treatment data. The response surface model predictions and the confidence intervals overlap at 10 and 16–18 d, but there is some disagreement at intermediate treatment durations, indicating that the model fails to account for all of the observed variation. Similarly, the regression analysis of the Back and Pemberton (1916) data also indicates a large degree of overdispersion about the response surface model (Table 2), suggesting that alternative models should be considered for the purpose of improved response surface modeling. The biological mechanism of larval mortality because of low temperature is not well understood, but if a critical physiological point exists (e.g., beyond which cell walls rapidly lose integrity), this might suggest using a discontinuous (e.g., splined) model form. Many discontinuous surface modeling approaches suffer from a distinctly *ad hoc* flavor, however, underscoring the fundamental importance of understanding the underlying biology. Powell (2002) presents a response surface model for heat treatment of solid wood to eliminate fungal pathogens using an approach that relaxes the parametric assumptions associated with traditional statistical methods and better reflects irregularities in the observed data. Applying such an approach to modeling the response of *C. capitata* larvae to cold treatment, however, could be fairly characterized as statistical overkill given the nature of the currently available data.

Discussion

Although the original work was conducted >80 yr ago, the conclusions drawn from the response surface model appear reasonably robust in comparison to more recent studies and surveillance data. Overall, the quantitative analysis suggests that within the range considered (32–36°F (0–2.2°C)), the duration of cold treatment may be more important than the nominal storage temperature in driving *C. capitata* larval survival to very low levels. Cold treatment data from recent studies are sparse. Given the vintage of some of the data, methodological inconsistencies, and consequences of new introductions, additional research is warranted, especially to verify the efficacy of low-temperature, short-duration treatments.

Although the broad coverage of the Back and Pemberton (1916) data makes it best suited for constructing a time–temperature response surface model for *C. capita* larval survival, for the purposes of revising the regulatory cold treatment schedule, elaborating a complete response surface and refining its fit are unnecessary. Instead, the efficacy of discrete time-temperature combinations may be evaluated independently. The evaluation data were analyzed assuming only that the uncertainty regarding the proportion of survivors can be characterized by the beta distribution, i.e., larval morality results from a binomial process with unknown but invariant p_s. Although alternative assumptions also could be explored (e.g., density-dependent p_s), focusing on a limited set of discrete time–temperature combinations permits us to relax or simply avoid the far more numerous statistical assumptions inherent to response surface methods (e.g., the assumption that the logit transformation (equation 1) is linearizing). This is of particular concern because predictions at the extremely low survival levels relevant to phytosanitary programs may be dominated not by the observed data but by the assumed statistical model form (e.g., heavy-tailed or light-tailed distribution). Given that more recent studies have illuminated some of the discrete time–temperature cold treatment combinations of primary concern, the greatest remaining uncertainty appears to be whether treatments of <14 d at temperatures in the 32–33°F (0.0–0.6°C) range will achieve the probit nine level of security. The analysis suggests, therefore, that the efficacy of low-temperature, short-duration treatments should be a primary focus of new data acquisition.

The practical significance of the Bayesian statistical methods used to analyze the evaluation data are that researchers need not conduct unfeasibly large trials to assess the performance of cold treatments resulting in very low levels of insect survival. Some may believe that only trials with a very large number of treated insects will permit statistical analysis of probit nine security levels. Consider, however, that for a trial with zero observed survivors to provide 95% confidence that probit nine level security has been achieved, the number of treated insects would have to exceed 90,000 (Couey and Chew 1986). Figure 5 presents the diminishing returns of increasing the number of test insects (with zero survivors) to confidence in achieving the probit nine level of security. The belief that such mega-trials are necessary can pose a strong deterrent to initiating needed research.

Although the "absence of evidence is not evidence of absence," the number of treated insects need not be impracticably large to permit informative analysis. The Bayesian statistical methods described above demonstrate that trials with zero survivors can provide useful information about the true underlying probability of survival at a given time-temperature combination. In such cases, it is more than intuitive that our

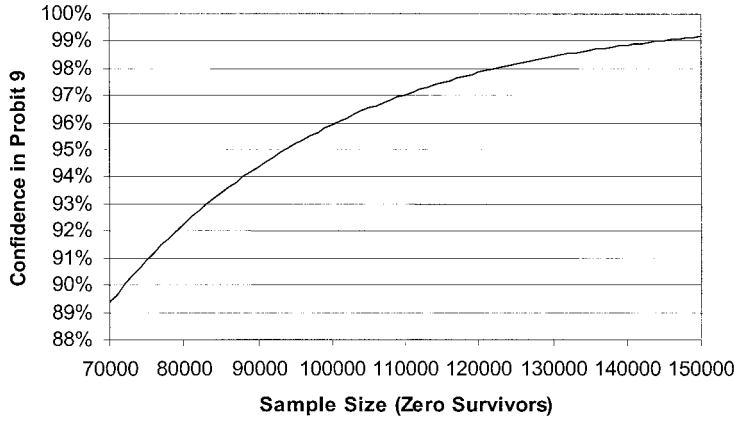

Fig. 5. Relationship between sample size and confidence in probit 9.

confidence that the probability of survival is small increases with sample size. Therefore, replicated trials of more manageable size offer a feasible alternative to an impracticably large mega-trial. Conveniently, as the cumulative sample size becomes large (e.g., $n = n_i > 10{,}000$), the statistical estimation process can be greatly simplified by noting that the beta distribution parameter estimates resulting from the laborious Bayesian procedure closely approximate those obtained by assuming a uniform (0,1) prior (implying that in the absence of information, all values of p_s between zero and one are considered equally likely) and a single, large trial (Vose 2000):

$$\hat{\alpha} = s + 1, \qquad [6]$$

$$\hat{\beta} = n - s + 1.$$

A legitimate concern arising from the use of Bayesian statistical methods is that the prior distribution is subjectively defined. In cases in which data are sparse, the Bayesian prior distribution may dominate the observed data in determining the resultant posterior distribution. As data accumulate, however, the influence of the prior distribution diminishes, and the empirical data come to statistically dominate the posterior distribution (Robert and Casella 1999).

Entirely separate from the question of whether the cold treatment attains the intended level of mortality is whether the probit nine level of security is either necessary or sufficient to maintain an acceptably low risk of establishment of new *C. capitata* populations outside the pest's current distribution. Given a large enough volume of infested fruit imports, even the probit nine level of security could be overwhelmed. However, attaining a greater level of security via treatment alone may be impracticable. Cold treatment, however, is not the only hurdle to clear. There are multiple sources of resistance to establishment of a new colony, which depends on pretreatment infestation levels, dynamics of escaping mortality and predation in a novel ecological community, synchronous emergence of adult male and female survivors, density-dependent probability of encountering a mate, spatially and temporally specific likelihood of encountering a suitable host for oviposition, and other factors. The determination of the appropriate level of phytosanitary protection in any specific case poses a serious policy challenge for risk managers who seek to balance a complex, uncertain, and unevenly distributed set of risks, costs, and benefits.

References Cited

Back, E. A., and C. E. Pemberton. 1916. Effect of cold-storage temperatures upon the Mediterranean fruit fly. J of Ag Res. 5: 657–666.

Brown, D., and P. Rothery. 1993. Models in biology: mathematics, statistics and computing. Wiley, NY.

Couey, H. M., and V. Chew. 1986. Confidence limits and sample size in quarantine research. J. Econ. Entomol. 79: 887–890.

De Lima, C.P.F., A. Jessup, and R. McLauchlan. 2002. Cold disinfestation of citrus using different temperature x time combinations. Horticulture Australia, South Perth, Australia. (Available at: APHIS Docket No. 02–023-4).

Evans, M., N. Hastings, and B. Peacock. 1993. Statistical distributions, 2nd ed. Wiley, NY.

Hill, A. R., C. J. Rigney, and A. N. Sproul. 1988. Cold storage of oranges as disinfestation treatment against the fruit flies *Dacus tyroni* (Froggatt) and *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae). J of Econ Entomol. 81: 257–260.

Jessup, A. J., C.P.F. De Lima, C. W. Hood, R. F. Sloggett, A. M. Harris, and M. Beckingham. 1993. Quarantine disinfestation of lemons against *Bactrocera tyroni* and *Ceratitis capitata* (Diptera: Tephritidae) using cold storage. J of Econ Entomol. 86: 798–802.

Mason, A. C., and O. C. McBride. 1934. Effect of low temperatures on the Mediterranean fruit fly in infested fruit. J of Econ Entomol. 27: 897–902.

Nel, R. G. 1936. The utilization of low temperatures in the sterilization of deciduous fruit infested with the immature stages of Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann). Union of South Africa Department of Agriculture and Forestry, Pretoria, South Africa.

Powell, M. 2002. A model for probabilistic assessment of phytosanitary risk reduction measures. Plant Disease. 85: 552–557.

Robert, C. P., and G. Casella. 1999. Monte carlo statistical methods. Springer, NY.

Santaballa, E., R. Laborda, and M. Cerda. 1999. Informe sobre tratamiento frigorifico de cuarentena contra *Ceratitis capitata* (Wied) para exportar mandarinas clementinas a Japon. Universidad Politecnica de Valencia, Valencia, Spain. (Available at: APHIS Docket No. 02–023-4).

Shipguide. com. 2002. The Electronic Shipping Guide. (http://www.shipguide.com).

Sproul, A. N. 1976. Disinfestation of Western Australian Granny Smith apples by cold treatment against the egg and larval stages of the Mediterranean fruit fly, Ceratitis capitata (Wiedemann). Austral J of Exp Ag and Animal Husbandry. 16: 280–285.

[USDA—APHIS] U.S. Department of Agriculture–Animal and Plant Health Inspection Service. 2002a. Evaluation of cold storage treatment against Mediterranean Fruit Fly, *Ceratitis capitata* (Wiedemann) (Diptera:Tephritidae). Oxford Plant Protection Center, Oxford, NC. (http://www.aphis.usda. gov/oa/clementine/)

[USDA—APHIS] U.S. Department of Agriculture–Animal and Plant Health Inspection Service. 2002b. Treatment manual. Plant Protection and Quarantine, Washington, DC. (http://www.aphis.usda. gov/ppq/manuals/pdf_files/TM. pdf)

[USDA—APHIS] U.S. Department of Agriculture–Animal and Plant Health Inspection Service. 2002c. Importation of clementines from Spain: proposed rule. Federal Register 67: 45922–45933.

[USDA—APHIS] U.S. Department of Agriculture–Animal and Plant Health Inspection Service. 2002d. Importation of clementines from Spain: final rule. Federal Register 67: 64702–64739.

[USDA—APHIS] U.S. Department of Agriculture–Animal and Plant Health Inspection Service. 2002e. Risk mitigation for Mediterranean fruit flies with special emphasis on risk reduction for commercial imports of clementines (several varieties of *Citrus reticulata*) from Spain. Center for Plant Health Science and Technology, Raleigh, NC. (http://www.aphis.usda. gov/oa/clementine/)

Vose, D. 2000. Risk analysis: a quantitative guide, 2nd ed. Wiley, Chichester, England.

Received for publication 5 July 2002; accepted 27 November 2002.